

文章编号:1001-9081(2013)09-2577-03

doi:10.11772/j.issn.1001-9081.2013.09.2577

基于运动估计与回溯自适应正交匹配追踪的视频压缩传感

庄燕滨^{1,2}, 桂源^{1*}, 肖贤建²

(1. 河海大学 计算机与信息学院, 南京 211100; 2. 常州工学院 计算机信息工程学院, 江苏 常州 213002)

(*通信作者电子邮箱 anglowe@yeah.net)

摘要:为了解决传统视频压缩传感方法中对视频逐帧单独重构所产生的图像模糊,将压缩传感理论与MPEG标准视频编码的相关技术相结合,提出了一种基于运动估计与运动补偿的视频压缩传感方法,以消除视频信号在空间和时域上的冗余。该方法在充分考虑视频序列时域相关性的同时,首先对视频图像进行前、后向和双向预测和补偿,然后采用回溯自适应正交匹配追踪(BAOMP)算法,对运动预测残差进行重构,最后实现当前帧的重构。实验结果表明,该方法较逐帧重构的视频图像质量有较大改善,且可获得更高的峰值信噪比。

关键词:压缩传感;视频压缩;运动估计;时域相关性;回溯自适应正交匹配追踪

中图分类号:TP317.4 文献标志码:A

Compressed video sensing method based on motion estimation and backtracking-based adaptive orthogonal matching pursuit

ZHUANG Yanbin^{1,2}, GUI Yuan^{1*}, XIAO Xianjian²

(1. College of Computer and Information Technology Engineering, Hohai University, Nanjing Jiangsu 211100, China;

2. School of Computer Information and Engineering, Changzhou Institute of Technology, Changzhou Jiangsu 213002, China)

Abstract: In order to remove the image blurring caused by reconstructing video frames independently frame by frame using traditional compressed video sensing method, this paper proposed a new approach to video compressed sensing based on motion estimation and motion compensation by combining the compressed sensing theory with related technology of MPEG standard video coding, so as to remove the spatial and temporal redundancy of video signal. This method took full account of the temporal correlations of video sequences and firstly compensated video frames using forward, backward and bidirectional prediction, then adopted the Backtracking-based Adaptive Orthogonal Matching Pursuit (BAOMP) algorithm to reconstruct the motion prediction residuals and finally reconstructed current frames. The experimental results indicate that the proposed method can gain a better video image quality compared with frame-by-frame reconstruction method and achieve a higher Peak Signal-to-Noise Ratio (PSNR).

Key words: compressed sensing; video compression; motion estimation; temporal correlation; Backtracking-based Adaptive Orthogonal Matching Pursuit (BAOMP)

0 引言

随着3G移动技术和视频监控技术的不断发展,如何提高压缩率、节省存储空间并高效地传输视频信号成为视频压缩领域的一个重要研究问题。考虑到多数信号是可压缩的,Donoho等在2006年提出了一种新的信号获取理论:压缩传感(Compressed Sensing, CS)^[1-2]理论,突破了传统Nyquist方法采样率不小于最高频率两倍的限制。该理论针对稀疏信号,合并了采样和压缩两个过程,直接将高维信号投影到低维空间,从而实现欠Nyquist采样^[3],有效地缓解了编码端的压力^[4]。因视频由多帧图像构成,本身含有大量空间冗余(静止图像行列相邻点的相关性)信息,是可以稀疏表示的,因此可以根据CS理论对视频帧进行压缩和重构。近年来,国内外很多学者都开始对视频的压缩传感方法进行研究,以求在保证视频重构质量的前提下降低视频采集端的代价。文献[4]将三维双数复数小波(3D Dual-Tree Complex Wavelet Transform, 3D DTCWT)作为视频的稀疏变换域,结合迭代收

缩(Iterative Shrinkage/Threshold, IST)算法实现了视频的压缩重构,得到了较好的结果,但对快速运动的视频表现不很理想。文献[5]提出了一种分布式视频压缩传感方法(Distributed Compressed Video Sensing, DCVS),其本质是利用压缩传感理论解决分布式视频压缩编码问题,实现了低复杂度编码,在实际应用(如:THz成像)中有很好的前景。而为了提高利用CS理论进行视频压缩重构的质量,本文提出了一种基于运动估计和回溯自适应匹配追踪的视频压缩传感算法。

本文考虑到视频中物体的运动具有时域上的相关性(连续帧的相似性),若直接对视频逐帧使用CS方法进行帧内编码,将出现因帧间抖动而产生的图像模糊,影响视觉质量。参考现有的视频压缩编码技术,如MPEG系列标准的视频压缩编码技术,其主要利用了运动估计和补偿来消除帧间的时域相关性。因此,本文将结合压缩传感理论和运动估计和补偿技术来实现视频的压缩传感^[6-7],以获取良好的视频压缩重构图像。

收稿日期:2013-03-29;修回日期:2013-05-06。

基金项目:江苏省科技厅工业支撑计划项目(BE2010072);常州市科技局国际合作项目(CZ20123006)。

作者简介:庄燕滨(1964-),男,江苏常州人,教授,主要研究方向:智能信息处理、视频图像处理、模式识别;桂源(1988-),女,云南临沧人,硕士研究生,主要研究方向:视频图像处理;肖贤建(1974-),男,江西宜丰人,副教授,博士,主要研究方向:无线传感网、视频图像处理。

文献 [8] 提出了一种回溯自适应正交匹配追踪(Backtracking-based Adaptive Orthogonal Matching Pursuit, BAOMP)算法,在传统的正交匹配追踪(Orthogonal Matching Pursuit, OMP)中结合了一种回溯算法。其本质是在每次迭代时检测之前原子的可靠性,然后删除不可靠原子,使迭代结果更优,且可实现稀疏度的自适应。这种算法在获得良好重构质量的同时继承了OMP算法低复杂度的优点。因此,为了得到更精确的解码结果,本文选用BAOMP算法作为重构算法,并实验对比了该算法和几种现有算法的重构效果。

1 CS 理论

CS 理论主要包括信号的稀疏表示、投影测量和信号重构三个方面^[9]。

长度为 N 的一维离散信号 x 可表示为:

$$x = \sum_{i=1}^N \theta_i \psi_i \quad \text{或} \quad x = \Psi \theta \quad (1)$$

其中: $N \times N$ 的正交基(稀疏基) $\Psi = [\psi_1 | \psi_2 | \dots | \psi_N]$, ψ_i 为 $N \times 1$ 的列向量,若 θ 只有 K ($K \ll N$) 个非零值,则为信号 x 的等价 K 稀疏表示, $\theta_i = \langle x, \psi_i \rangle = \psi_i^T x$ 。

为了降低采样数,CS 的采样测量过程将信号 x 投影到更低维的空间。假设已知某一测量矩阵 $\Phi \in \mathbb{R}^{M \times N}$ ($M \ll N$), $y \in \mathbb{R}^M$ 为信号 x 在该矩阵下的线性投影测量值,则

$$y = \Phi x \quad (2)$$

根据式(1)得稀疏信号的投影测量为:

$$y = \Phi x = \Phi \Psi \theta = \tilde{\Phi} \theta \quad (3)$$

因 y 的维数远远低于 θ 的维数,方程(3)是一个 NP-hard 问题,可以由测量值 y 通过求解 l_0 范数最优化问题重构信号:

$$\tilde{\theta} = \operatorname{argmin} \|\theta\|_0$$

$$\text{s. t. } \tilde{\Phi} \tilde{\theta} = y \quad (4)$$

其中: $\|\cdot\|_0$ 为向量的 l_0 范数,表示向量 x 中非零元素的个数。但前提是测量次数 M (即 y 的维数) 满足 $M = O(K \log(\frac{N}{K}))$, 并且矩阵 $\tilde{\Phi} = \Phi \Psi$ 满足约束等距性(Restricted Isometry Property, RIP)条件^[10], 等价条件是测量矩阵 Φ 必须完全和稀疏基 Ψ 完全不相干。

通过解式(4)得到信号 x 的稀疏表示 $\tilde{\theta}$, 再将结果代入式(1)中求解出 \tilde{x} , \tilde{x} 即为重构的原始信号。

2 视频压缩传感算法

2.1 运动估计与补偿模型

为了改善利用 CS 方法对视频帧逐帧重构的图像质量并考虑到视频帧的时域相关性,本文使用了一种基于视频帧运动估计与补偿的 3D 模型,如图 1 所示。

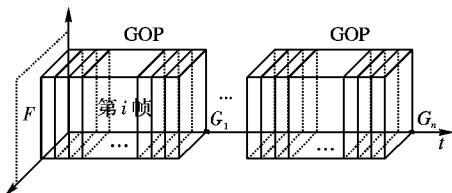
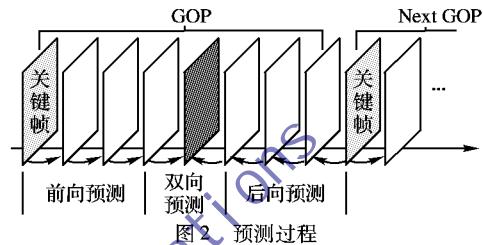


图 1 基于视频帧运动估计与补偿的 3D 模型

其本质是将视频帧按编码顺序分为等分的视频帧组(Group Of Pictures, GOP),每个视频帧组可看作一个 3D 的编码模型,每个 GOP 的第一帧作为关键帧,其余为非关键帧。

在运动补偿预测阶段采用前向、后向和双向预测相结合的方法以获得准确的补偿预测结果。假设一个 GOP 含 8 个视频帧,则预测过程如图 2 所示。运动估计算法采用一种扩展自适应范围搜索(Extended Adaptive Range Search, EARS)快速算法^[11],该算法能根据视频本身的运动特性,自适应地确定搜索范围,使绝大部分最佳匹配点都落于该范围内,在保证峰值信噪比(Peak Signal-to-Noise Ratio, PSNR)方面性能的同时,避免了给定搜索窗时可能造成的不必要的浪费(对于小运动序列)或遗落最佳匹配点(对于大运动序列),且经实验证明,该算法在平均速度方面是全搜索算法的 280 多倍,是 3 步法的 2.8 倍,明显快于新 3 步算法和 4 步算法等经典搜索算法,大大节省了搜索时间。



2.2 算法描述

本文结合了压缩传感算法和运动估计与补偿相关算法,提出了一种视频压缩传感算法。为了降低对视频进行全局采样所产生的高复杂度,首先,对视频帧进行逐帧的独立随机采样^[12],得到各帧随机测量值 Y_n ,其中,对每个 GOP 关键帧的测量以高于非关键帧的测量比进行采样。然后,对关键帧和非关键帧分别以不同的方式进行重构:对关键帧 F_{key} ,利用给定的稀疏基和重构算法直接重构,得到 \tilde{F}_{key} ,以此作为下一帧的参考帧 F_{ref} ;而对于非关键帧的每个当前帧 F_n 采用前向、后向和双向预测的方式与参考帧(上一帧)进行补偿,得到当前帧的补偿帧 F_{c_n} ,当前帧与补偿帧的随机测量之差为当前帧预测残差的随机测量 Y_{r_n} ,即 $Y_n - Y_{c_n} = Y_{r_n}$,通过重构即可得到当前帧的残差帧 F_{r_n} 。又因 $Y_n = \Phi F_n = Y_{c_n} + Y_{r_n} = \Phi F_{c_n} + \Phi F_{r_n}$,最终得到 $F_n = F_{c_n} + F_{r_n}$,因此通过运动估计与补偿以及对运动预测残差的重构可获得当前帧的重构帧 \tilde{F}_n 。

2.3 解码算法

本文在解码阶段采用了回溯自适应正交匹配追踪(BAOMP)算法^[8]重构图像。算法通过二次筛选的方法选出可靠的原子,删除不可靠原子。

输入
 $\Phi: M \times N$ 测量矩阵;
 $y: M$ 维测量值;
 ε_1 : 预置原子添加常数阈值, $\varepsilon_1 \in [0, 1]$;
 ε_2 : 预置原子删除常数阈值, $\varepsilon_2 \in [0, 1]$;
 δ : 停止迭代的收敛阈值;
 n_{max} : 最大迭代次数。
初始化 初始解 $x_0 = \mathbf{0}$; 初始残差 $r_0 = y$; 估计支撑集 $A_0 = \emptyset$; 候补集 $\Pi_0 = \emptyset$; 删除集 $\Omega_0 = \emptyset$ 。
循环 令迭代次数 $i = 1$ 。
1) 预筛选。找到满足下式的原子,并添加进候补集 Π_i 中:
 $| \langle r_{i-1}, \varphi_{\Pi_i} \rangle | \geq \varepsilon_1 \cdot \max | \langle r_{i-1}, \varphi_j \rangle |$

其中: $j \in \{1, 2, \dots, N\}$, $|H_i| \leq M - |\Lambda|$ 。

2) 二次筛选。计算 $\mathbf{x}_i^{A \cup H_i} = (\Phi^{A \cup H_i})^T \cdot \mathbf{y}$, 选择原子删除集 Ω_i , 满足 $|\mathbf{x}_i^{A \cup H_i}| < \varepsilon_2 \cdot \max |\mathbf{x}_i^{H_i}|$ 。

3) 更新。更新 $\Lambda, \Lambda_i = \{\Lambda_{i-1} \cup H_i\} / \Omega_i$, 系数 $\mathbf{x}_i^A = (\Phi^A)^T \cdot \mathbf{y}$, 残差 $\mathbf{r}_i = \mathbf{y} - \Phi^A \mathbf{x}_i^A$ 。

4) 如果 $\|\mathbf{r}_i\|_2 < \delta$ 或 $i = n_{\max}$, 则退出迭代过程;否则,令 $i = i + 1$, 执行步骤1), 直到迭代收敛, $i = n_{\max}$ 退出循环。

输出 估计支撑集和 $\mathbf{x}_{\max}^A = (\Phi^A)^T \cdot \mathbf{y}$ 。

为了说明BAOMP算法的重构效果,本文以静止图像为

处理对象,将该算法与几种不同的重构算法进行了对比。首先对图像进行离散小波变换(Discrete Wavelet Transform, DWT),再以不同的测量比进行随机采样,用几种现有的算法重构 256×256 的静止图像 Scene 和 Town, 其中给定 $\varepsilon_1 = 0.5$, $\varepsilon_2 = 0.7$, $n_{\max} = M$, $\delta = 10^{-6}$ 。表1为BAOMP算法和压缩采样匹配追踪(Compressive Sampling Matching Pursuit, CoSaMP)^[13]、正交匹配追踪(OMP)^[14]、子空间追踪(Subspace Pursuit, SP)^[15]、迭代加权最小二乘(Iteratively Reweighted Least Square, IRLS)^[16]重构算法的PSNR值比较。

表1 不同算法的重构效果(平均PSNR值)比较

测量比 (M/N)	Scene					Town					dB
	CoSaMP	OMP	SP	IRLS	BAOMP	CoSaMP	OMP	SP	IRLS	BAOMP	
0.1	8.159	8.256	8.586	9.547	8.683	6.910	6.080	6.134	7.804	7.324	
0.2	13.392	13.600	13.856	13.849	14.016	9.688	9.502	10.121	11.437	10.019	
0.3	18.415	20.666	20.524	20.865	20.732	18.254	18.211	18.714	17.341	18.837	
0.4	23.652	25.058	24.979	24.969	24.880	21.049	21.049	21.402	21.730	22.026	
0.5	27.798	27.579	27.892	28.797	29.024	22.294	22.535	23.220	24.744	24.126	
0.6	29.962	30.534	30.222	32.416	32.568	25.404	25.007	25.736	27.303	28.065	

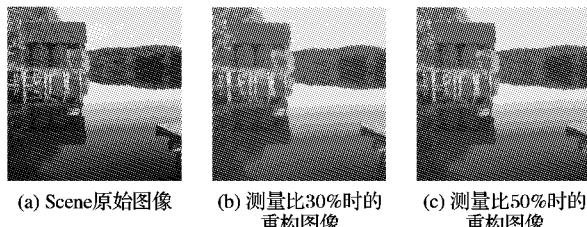


图3 Scene原始图像和重构图像

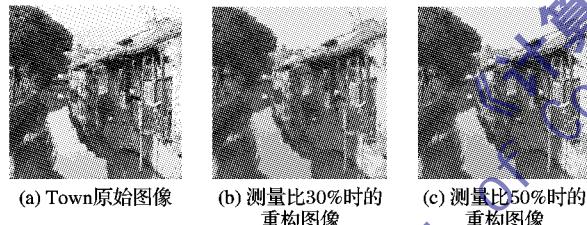


图4 Town原始图像和重构图像

3 实验结果及分析

实验选取 Robot 序列的一部分($256 \times 256, 64$ 帧)作为测试序列,比较了采用本文算法和对视频帧逐帧独立使用CS算法图像效果,如图5所示。实验将每个视频帧组GOP设为8,以测量比 $r_1 = 0.8$ 和 $r_2 = 0.3$ 分别对每个GOP的关键帧和非关键帧对视频帧进行独立随机测量,选择具有移位不变性和良好方向选择性的双树复数小波(Dual-Tree Complex Wavelet Transform, DT-CWT)作为稀疏基,BAOMP算法重构关键帧及帧间预测残差。为了降低运算复杂度,在解码时采用EARS快速搜索算法,选用 16×16 块,以绝对误差和(Sum of Absolute Difference, SAD)作为块匹配准则,以前向、后向和双向补偿预测方式获取补偿帧。在逐帧重构时也同样选用随机测量矩阵和DT-CWT稀疏基,测量比设为 $r = 0.3$ 。

由图5可见,逐帧使用CS方法所获得的图像尤其是运动物体的图像发生了模糊甚至丢失某些细节,这是由于此方法没有充分考虑具有运动物体的视频帧在时域上的相关性,因而产生帧间的抖动,出现模糊。而本文算法不仅通过在帧内使用CS方法消除了帧内空域相关性,而且结合运动估计和补偿方法去除了帧间时域相关性。通过比较可知,使用本文算法获得的图像无论从画质还是细节都比逐帧CS重构方法

有了很大的改善,且对应的每一帧的PSNR值分别提高了2.21 dB、2.47 dB和1.63 dB。

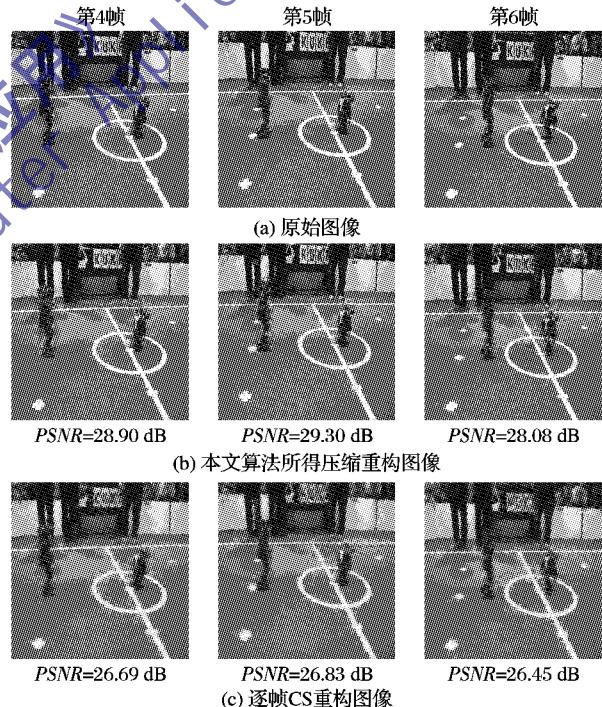


图5 Robot视频序列本文算法、逐帧CS重构的重构帧对比

4 结语

本文提出了基于GOP的3D模型,对视频逐帧随机采样后,首先以DT-CWT稀疏基和BAOMP算法重构关键帧,以去除帧内数据的相关性,然后通过前向、后向和双向预测,以EARS自适应搜索快速算法补偿当前帧,最后通过对预测残差的BAOMP重构从而得到当前帧的重构图像。算法结合运动估计与补偿技术去除了帧间相关性,实现了对视频的压缩传感。实验证明,本文算法具有可行性,且获得了较好的重构结果。但在数据量方面,尽管通过压缩传感去除了帧内冗余,但测量结果数据量依然较大,若考虑到实用性,还需进一步对测量结果进行压缩。

(下转第2587页)

表1 跟踪中心坐标偏差统计及平均跟踪速度对比

实验	算法	目标的跟踪中心坐标与手动标注中心坐标的偏差绝对值/像素						平均跟踪速度/ (frame · s ⁻¹)
		X-max	X-min	X-mean	Y-max	Y-min	Y-mean	
实验1	文献[5]算法	202	0	74.3	67	0	35.40	12.0
	文献[8]算法	178	0	55.6	46	0	9.60	9.6
	文献[10]算法	156	0	67.2	62	0	43.30	17.4
	本文算法	6	0	2.2	5	0	1.60	16.5
实验2	文献[5]算法	24	0	10.3	53	0	23.20	15.1
	文献[8]算法	22	0	12.1	18	0	9.30	12.3
	文献[10]算法	14	0	8.7	10	0	6.80	23.6
	本文算法	4	0	1.5	3	0	1.40	21.0

5 结语

通过大量分析前人的优秀研究成果,本文将目标的局部背景信息引入目标描述,减少背景干扰;其次在多特征融合跟踪过程中采用多准则的融合判定方式,有效地提高特征权值计算的准确性;最后嵌入到 Mean-shift 跟踪框架下来进行目标定位跟踪。在不同实验环境下表明:本文算法在背景复杂、光照变化和相似物干扰情况下都能较好地持续稳定跟踪,具有较好的跟踪鲁棒性。

参考文献:

- [1] 侯志强, 韩崇昭. 视觉跟踪技术综述[J]. 自动化学报, 2006, 32(4): 603–617.
- [2] COMANINI D, MEER P. Mean-shift: a robust approach toward feature space analysis [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(5): 603–619.
- [3] 王永忠, 梁彦, 赵春晖, 等. 基于多特征自适应融合的核跟踪方法[J]. 自动化学报, 2008, 34(4): 393–399.
- [4] 顾鑫, 王海涛, 汪凌峰, 等. 基于不确定度量的多特征融合跟踪[J]. 自动化学报, 2011, 37(5): 551–559.
- [5] 郑玉凤, 马秀荣, 赵晓琳, 等. 基于颜色和边缘特征的均值迁移目

标跟踪算法[J]. 光电子·激光, 2011, 22(8): 1232–1235.

- [6] 刘晴, 唐林波, 赵保军. 基于自适应多特征融合的均值迁移红外目标跟踪[J]. 电子与信息学报, 2012, 34(5): 1138–1141.
- [7] 袁广林, 薛模根, 韩裕生, 等. 基于自适应多特征融合的 mean shift 目标跟踪[J]. 计算机研究与发展, 2010, 47(9): 1663–1671.
- [8] COLLINS R T, LIU Y, LEORDEANU M. Online selection of discriminative tracking features[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(10): 1631–1643.
- [9] COMANICIU D, RAMESH V, MEER P. Kernel-based object tracking[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2003, 25(5): 564–577.
- [10] NING J, ZHANG L, ZHANG D, et al. Robust mean shift tracking with corrected background-weighted histogram[J]. IET Computer Vision, 2012, 6(1): 62–69.
- [11] WENG S K, KUO C M, TU S K. Video object tracking using adaptive Kalman filter[J]. Journal of Visual Communication and Image Representation, 2006, 17(6): 1190–1208.
- [12] 刘海燕, 杨昌玉, 刘春玲, 等. 基于梯度特征和颜色特征的运动目标跟踪算法[J]. 计算机应用, 2012, 32(5): 1265–1268.

(上接第 2579 页)

参考文献:

- [1] DONOHO D L. Compressed sensing[J]. IEEE Transactions on Information Theory, 2006, 52(4): 1289–1306.
- [2] CANDES E J. Compressive sampling[E]// Proceedings of the 2006 International Congress of Mathematicians. Piscataway, NJ: IEEE Press, 2006: 1433–1452.
- [3] CHANG H, ZHAO G, SHI H, et al. Polynomial rooting based frequency estimation with sub-Nyquist sampling[J]. Journal of Aeronautics, 2011, 32(1): 205–209.
- [4] 苏晓园, 练秋生. 基于迭代收缩法和三维复数小波的视频压缩传感重构[J]. 光电工程, 2010, 37(2): 108–112.
- [5] DO T T, CHEN Y, NGUYEN D T, et al. Distributed compressed video sensing[C]// CISS 2009: Proceedings of the 43rd Annual Conference on Information Sciences and Systems. Piscataway, NJ: IEEE Press, 2009: 1393–1396.
- [6] JUNG H, YE J C. Motion estimated and compensated compressed sensing dynamic magnetic resonance imaging: what we can learn from video compression techniques[J]. International Journal of Imaging Systems and Technology, 2010, 20(2): 81–98.
- [7] TROCAN M, TRAMEL E W, FOWLER J E, et al. Compressed-sensing recovery of multiview image and video sequences using signal prediction[J]. Multimedia Tools and Applications, 2013, 66(187): 1–27.
- [8] HUANG H L, ANAMITRA M. Backtracking-based matching pursuit method for sparse signal reconstruction[J]. IEEE Signal Processing Letters, 2011, 18(7): 391–394.
- [9] BARANIUK R G. Compressive sensing [lecture notes][J]. IEEE Signal Processing Magazine, 2007, 24(4): 118–121.
- [10] CANDES E J. The restricted isometry property and its implications for compressed sensing[J]. Comptes Rendus Mathématique, 2008, 346(9): 589–592.
- [11] 李翔, 吴国威. 一种适用于 H.264 的基于自适应搜索范围的快速运动估计算法[J]. 中国图象图形学报: A 辑, 2004, 9(4): 471–476.
- [12] MICHAEL W, JASON L, MARCO D, et al. Compressive imaging for video representation and coding[EB/OL].[2012-10-10]. <http://inside.mines.edu/~mwakin/papers/pcs-camera.pdf>.
- [13] NEEDELL D, TROPP J A. CoSaMP: Iterative signal recovery from incomplete and inaccurate samples[J]. Applied and Computational Harmonic Analysis, 2009, 26(3): 301–321.
- [14] TROPP J A, GILBERT A C. Signal recovery from random measurements via orthogonal matching pursuit[J]. IEEE Transactions on Information Theory, 2007, 53(12): 4655–4666.
- [15] WEI D, OLGICA M. Subspace pursuit for compressive sensing signal reconstruction[J]. IEEE Transactions on Information Theory, 2009, 55(5): 2230–2249.
- [16] RICK C, WOTAO Y. Iteratively reweighted algorithms for compressive sensing[C]// ICASSP 2008: Proceedings of the 2008 IEEE International Conference on Acoustics, Speech and Signal Processing. Piscataway, NJ: IEEE Press, 2008: 3869–3872.